# $H \to c\bar{c}$ search at LHCb and future upgrades

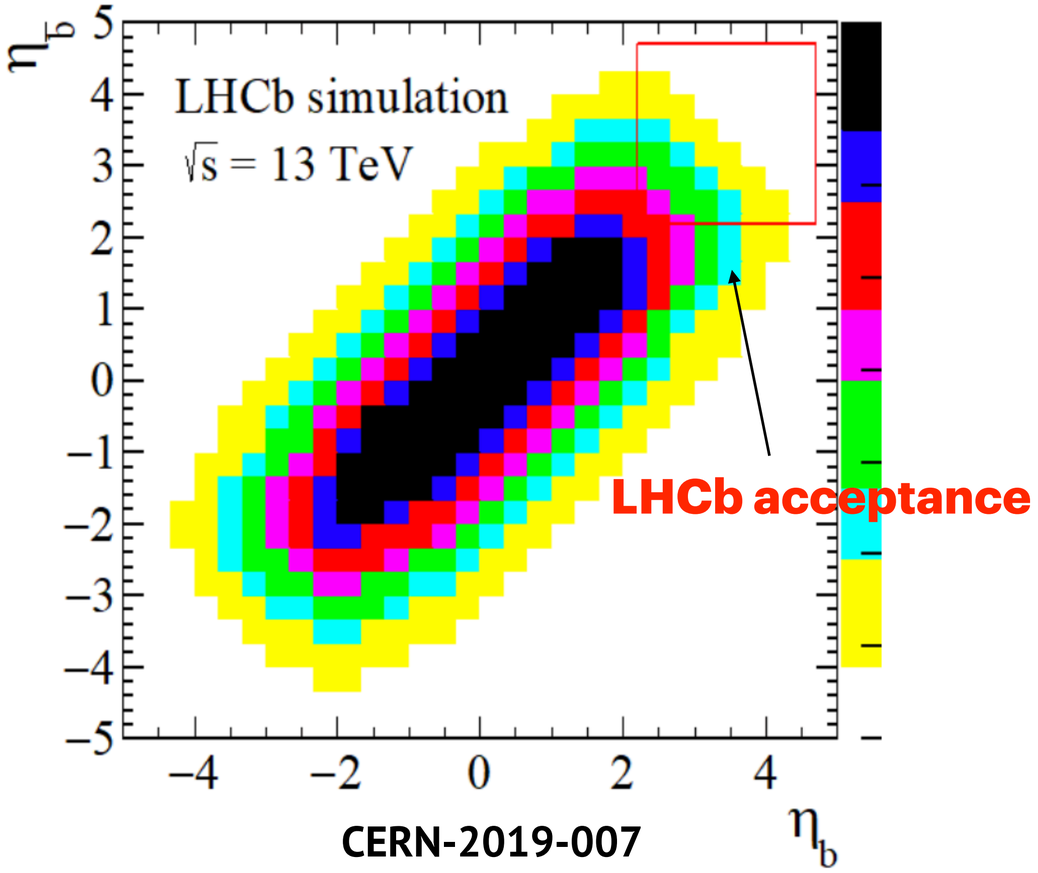
Lorenzo Sestini - INFN Padova

Snowmass EF01 Working Group meeting - 14/9/2021

# Why Higgs physics at LHCb?

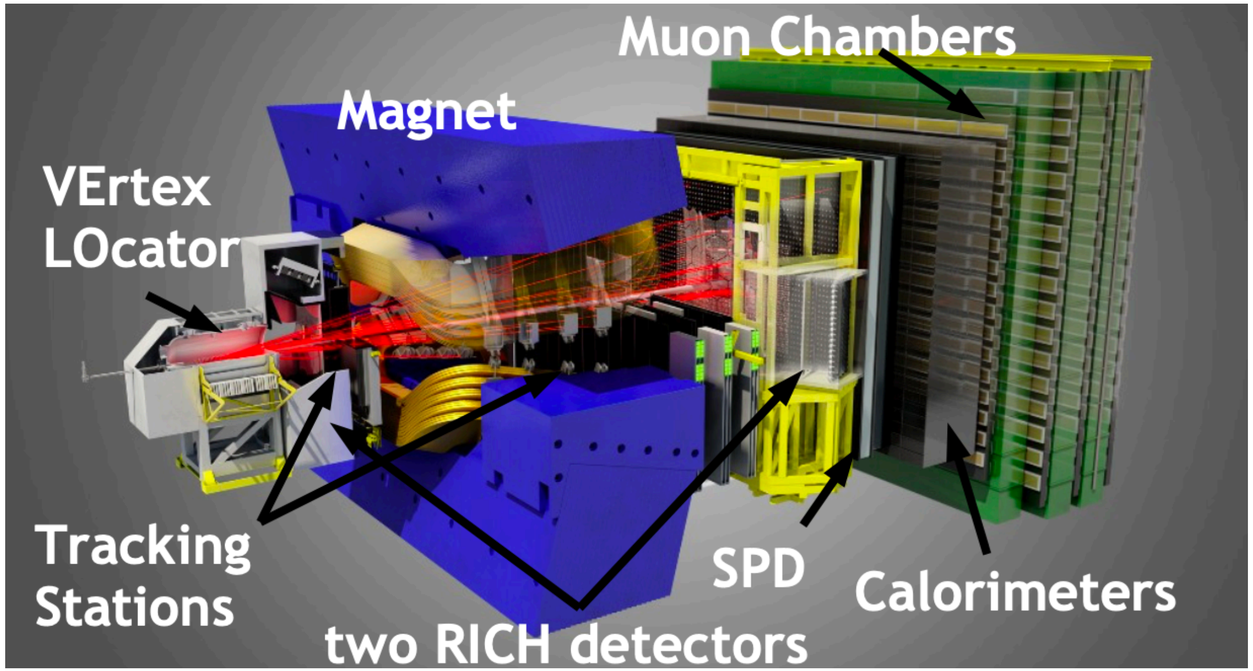
- At first sight **LHCb** may not be seen as a detector for Higgs physics:
  - Reduced acceptance with respect to ATLAS and CMS;
  - Lower luminosity due to leveling.
- But there are also **strong points**:
  - Lower pile-up means cleaner events;
  - Excellent secondary vertex reconstruction performance is a plus for b- and c-jets tagging.

**$H \to b\bar{b}$ pseudorapidity distribution**

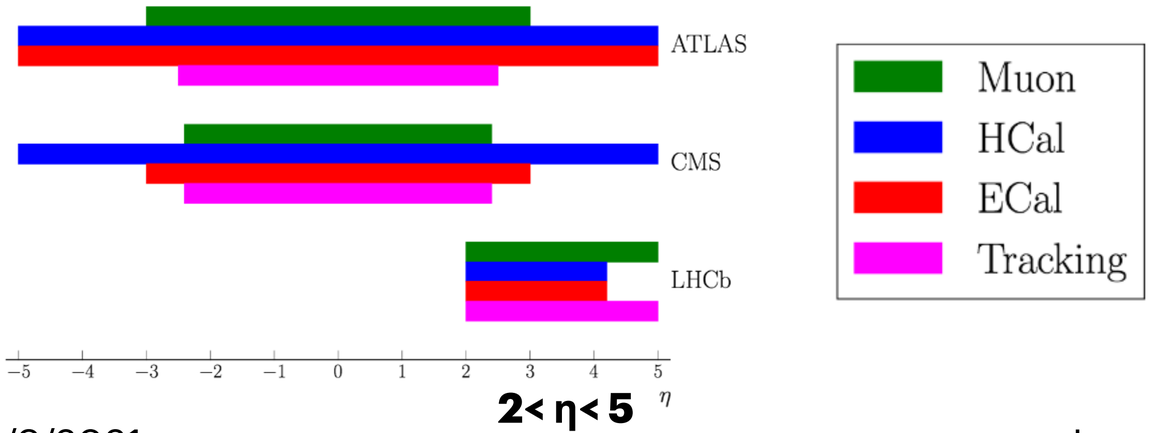


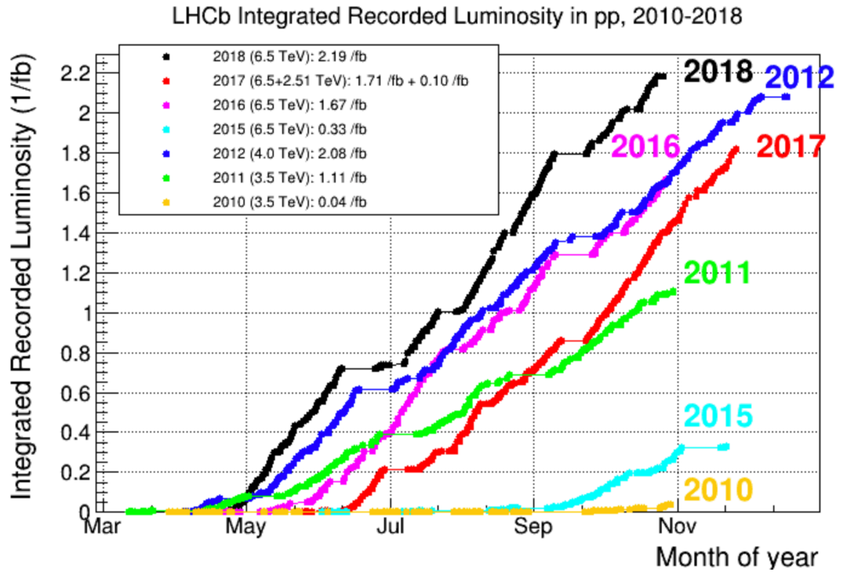
# LHCb detector

**JINST 3 (2008) S08005**



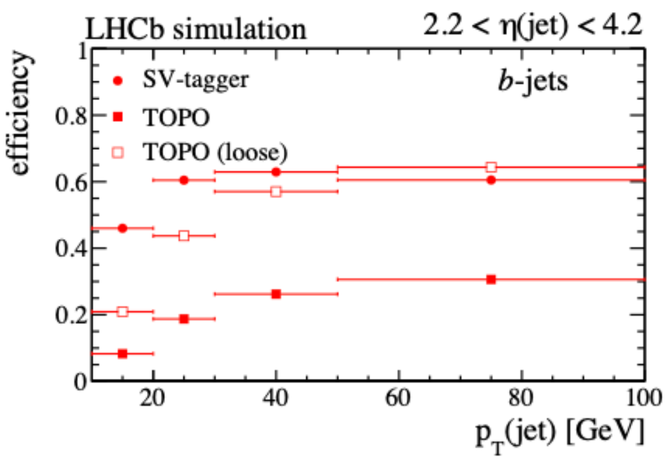

- LHCb, **originally designed for b- and c-hadron physics**, is now considered a **general purpose forward detector**.
- LHCb is a spectrometer in the forward region.
- Track momentum resolution: 0.4% at 5 GeV and 0.6% at 100 GeV.
- **Excellent vertex reconstruction system**: impact parameter resolution of $(15 + 29/p_T)$ μm, $p_T$ in GeV.
- Muon ID efficiency: 97% with 1-3% $\mu \rightarrow \pi$ mis-identification.
- Electron reconstruction with bremsstrahlung recovery.

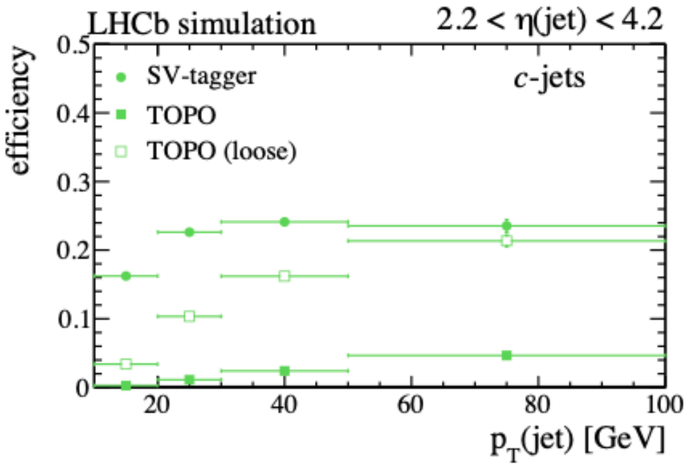


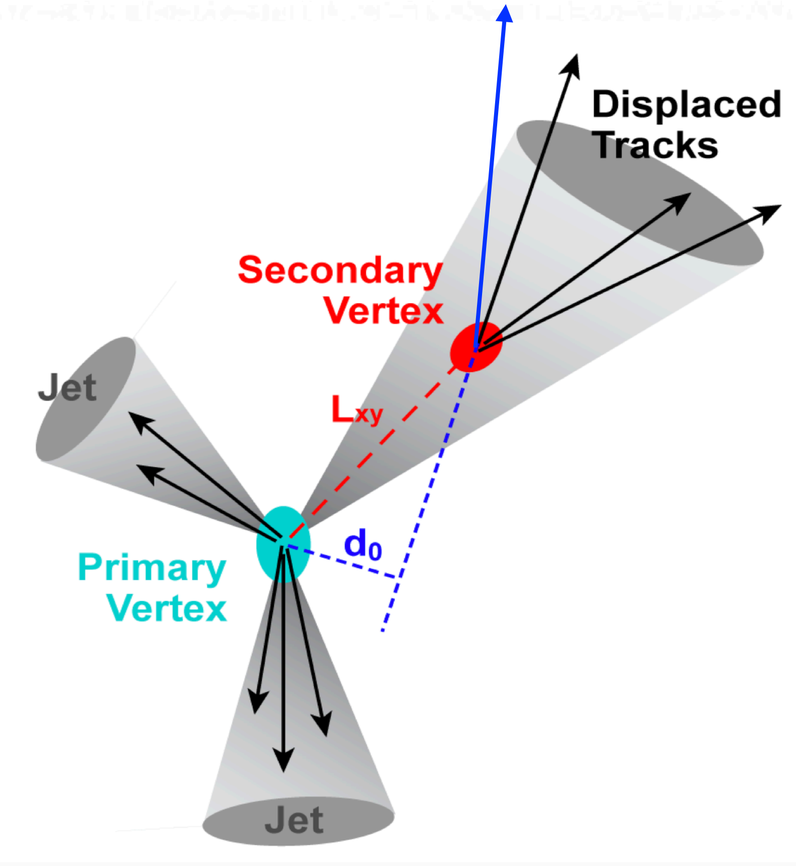



# Jet tagging at LHCb

- The jet tagging system takes advantage of LHCb features → **precise vertex reconstruction.**
- Secondary vertices are reconstructed **by using all the tracks in the events, not only those that belong to the jet**.
- Then a jet is tagged if a secondary vertex is found within distance of $\Delta R < 0.5$ with respect to the jet axis.
- In Run 1/2 t**he efficiency for tagging a b-(c-) jets is of about 60% (25%)**, **with a light jet misidentification below 0.1%**.




**JINST 10 (2015) P06013**

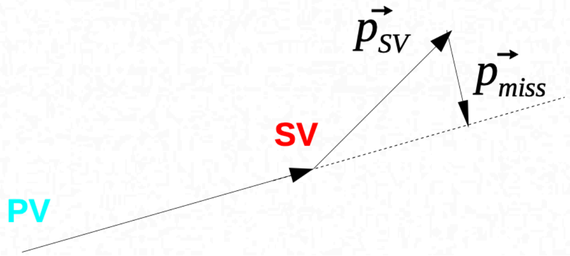
# Jet tagging at LHCb

- Two **Boosted Decision Trees** (BDT) are used to identify b and c jets.

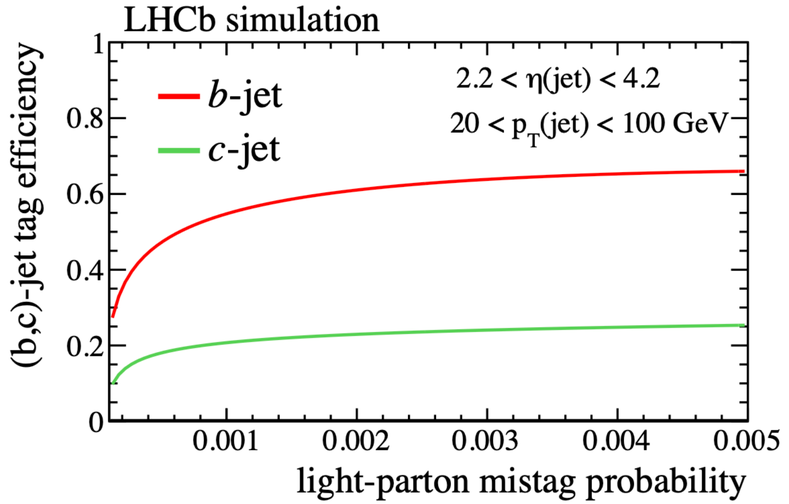
**BDT(bc|udsg)**
To separate **heavy flavour** jets from **light** jets

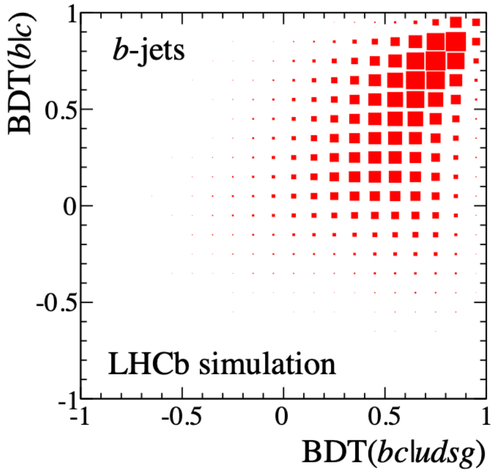
**BDT(b|c)**
To separate **b-jets** from **c-jets**

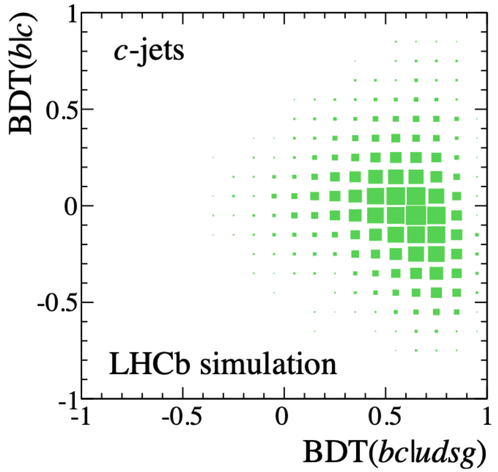
- **Some observables in input to the BDTs:**
  - **SV mass**
  - **Fraction of jet $p_T$ taken by the SV**
  - **Flight distance $\chi^2$**
  - **SV corrected mass**

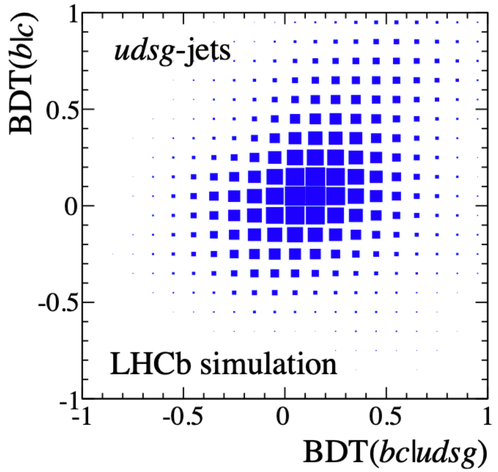



$$M_{corr} = \sqrt{M_{SV}^2 + p_{miss}^2} + p_{miss}$$



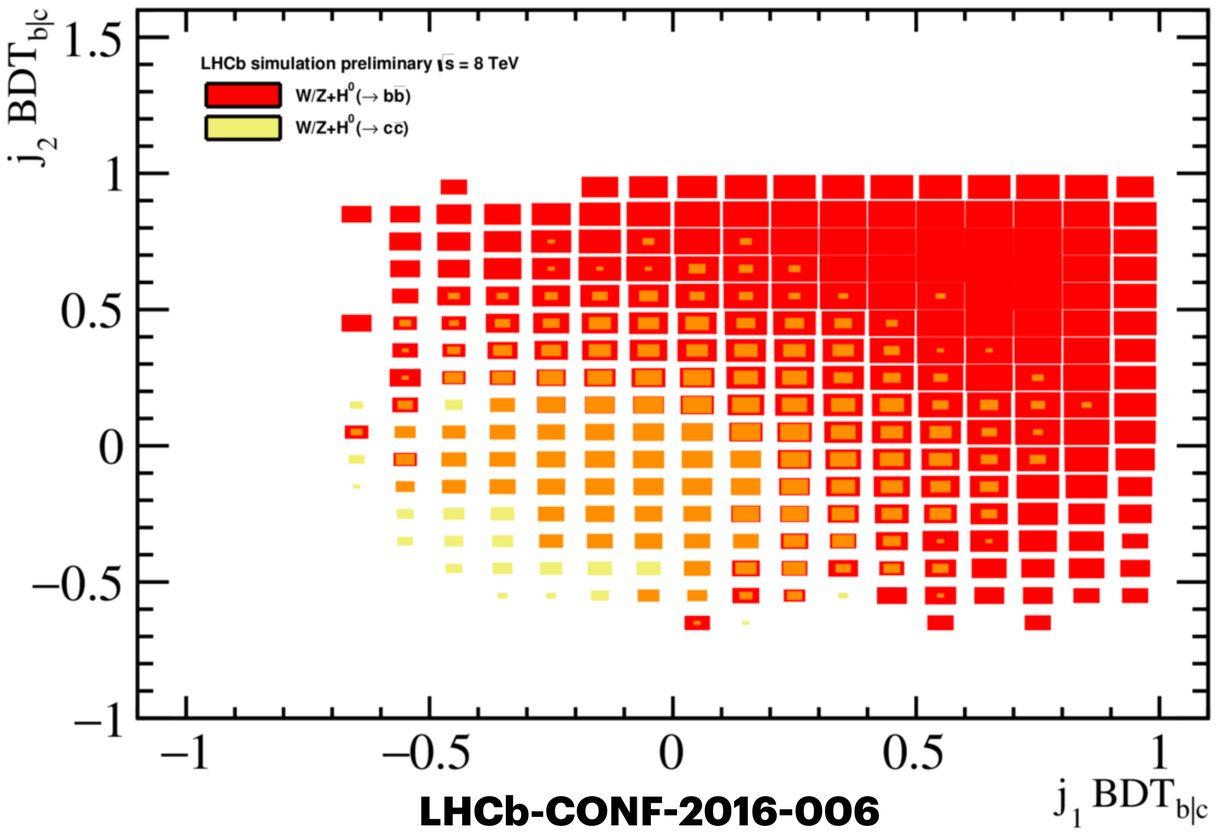



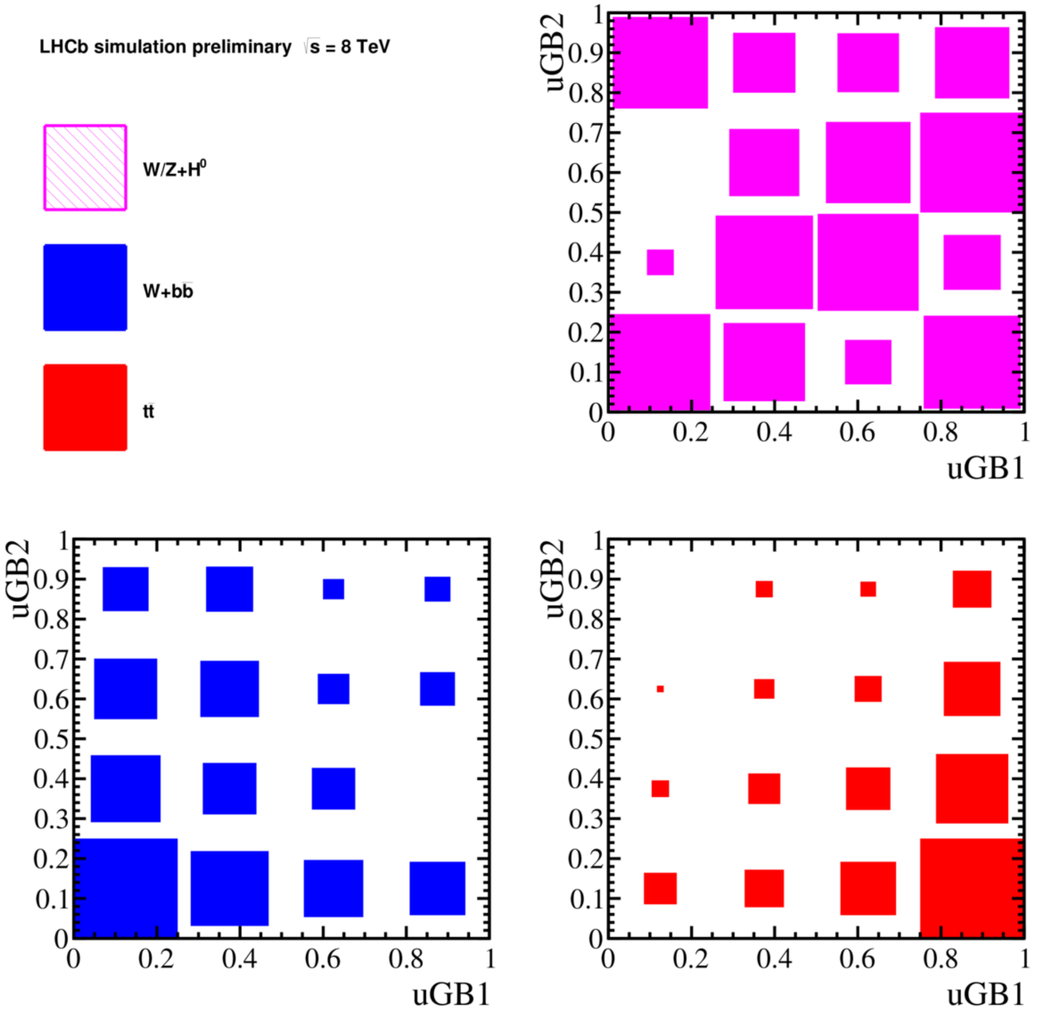
# H→ cc̄ search in Run 1: associated production

- LHCb searched for **W/Z + H (→cc̄)** in Run 1 data (**2 fb$^{-1}$ at 8 TeV**), by looking at the dijet+lepton (muon or electron) final state.
- The two jets are required to have $p_T$ >20 GeV and SV-tagged.
- One of the key aspects is that **H→bb̄ is an irreducible background for H→cc̄**.
- We can use the **BDT(b vs c)** to separate the two Higgs contributions.
- The optimal cut on BDT(b vs c) on both jets **removes 90% of H→bb̄ while keeping 62% of H→cc̄.**



# H→ cc̄ search in Run 1: associated production

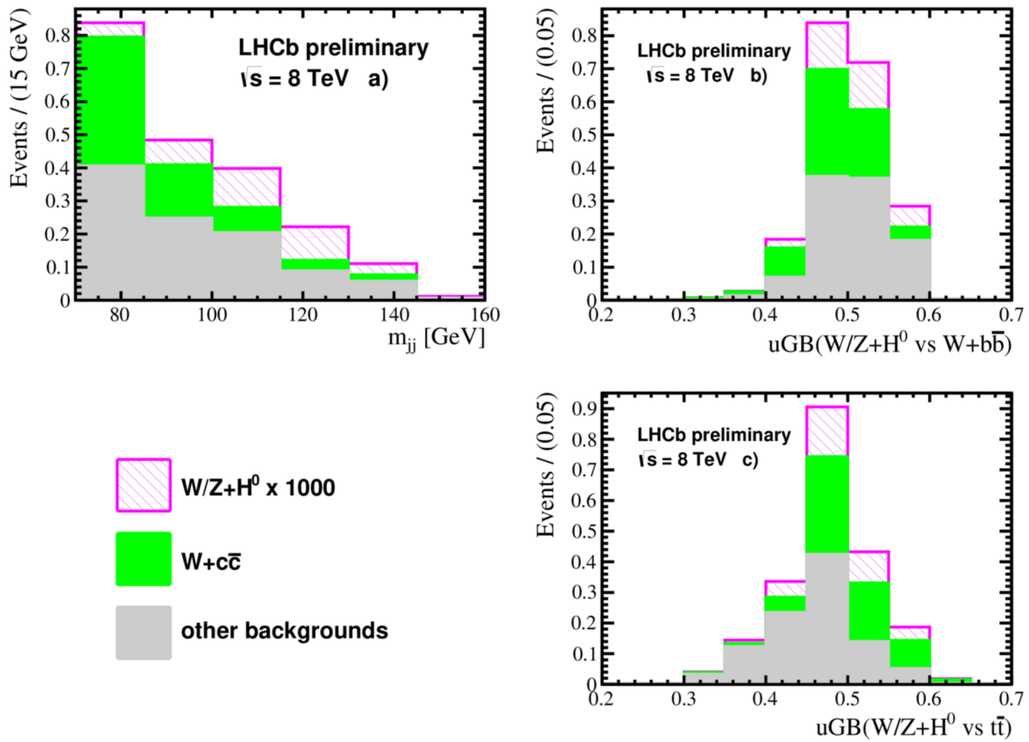
- The other backgrounds are **W+bb̄/cc̄ and top.**
- A multivariate method (**uniform gradient boost, uGB**) has been used to separate the backgrounds from the signal.
- Inputs are 12 kinematic variables of jets and leptons.
- **Two uGBs are trained in order to be uncorrelated with the dijet invariant mass**.
- A transformation is applied in order to make the distribution flat for the Higgs.



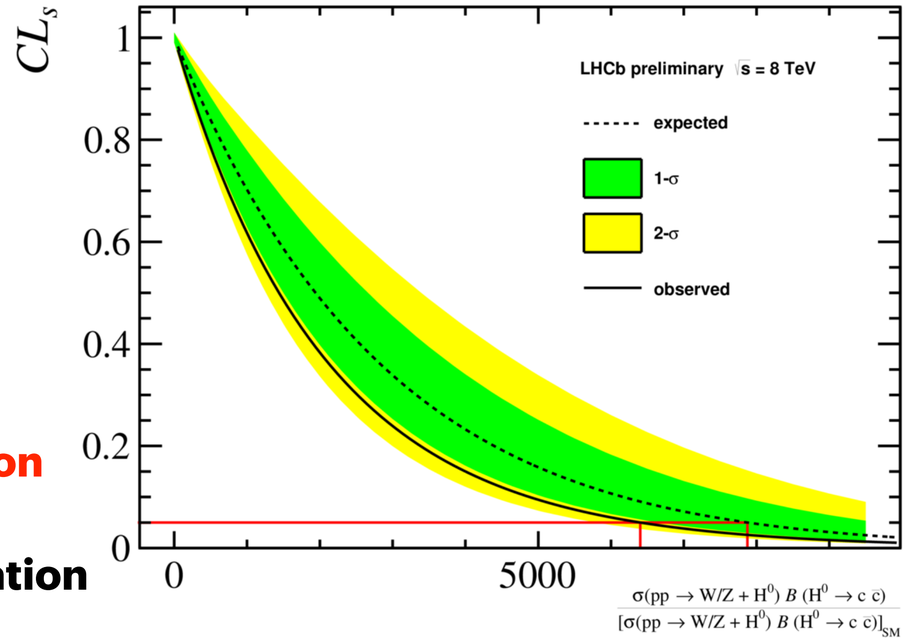

**LHCb-CONF-2016-006**

# H→ cc̄ search in Run 1: associated production

**LHCb-CONF-2016-006**




- **Of about 1.7 events are expected but 0 are observed.**
- The upper limit is set by using the dijet invariant mass and the two uGBs distributions.



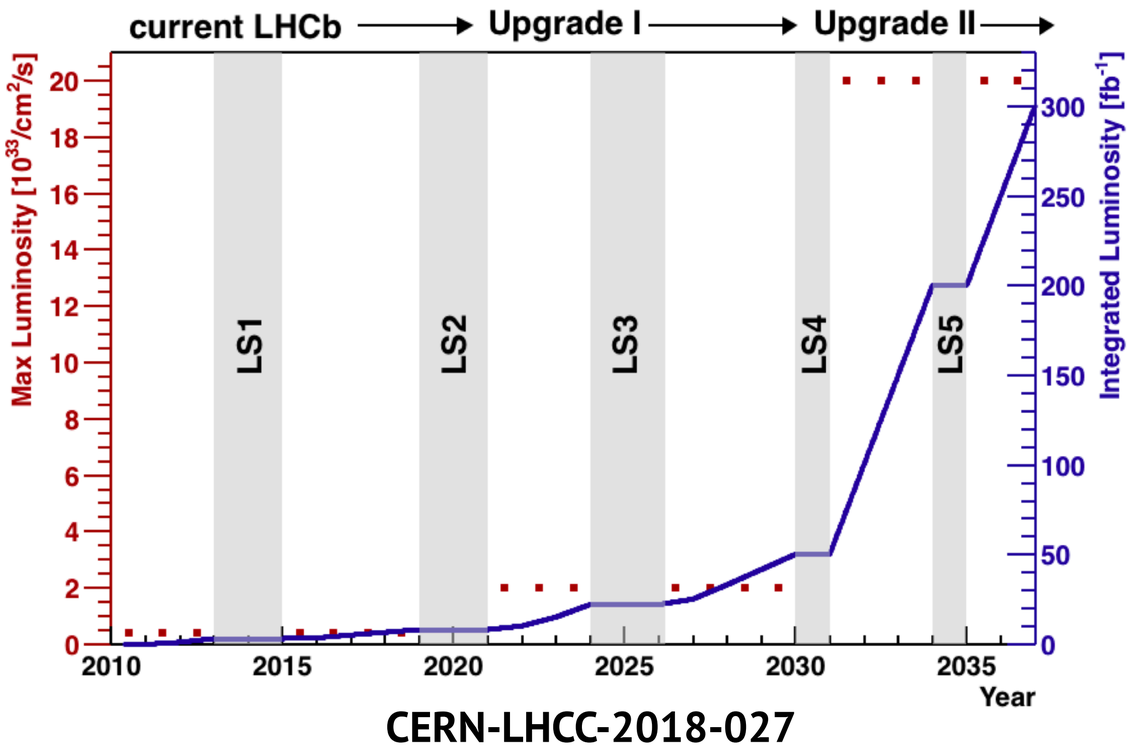

**The upper limit on σ x BR(H→cc̄) is 6400 times the SM expectation**

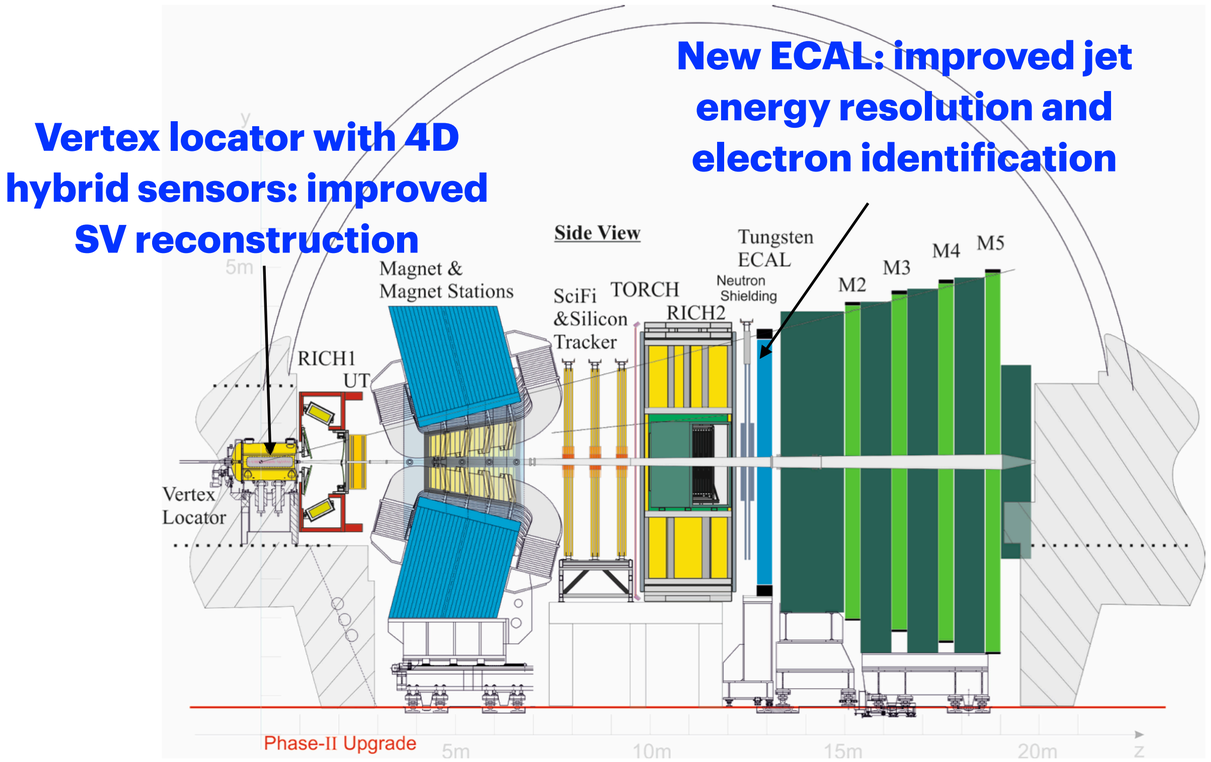
**The upper limit on the charm Yukawa coupling is $y^c < 80\ y^c_{SM}$**

**Cannot appear so good, but it is the starting point for our extrapolation**

# LHCb at HL-LHC

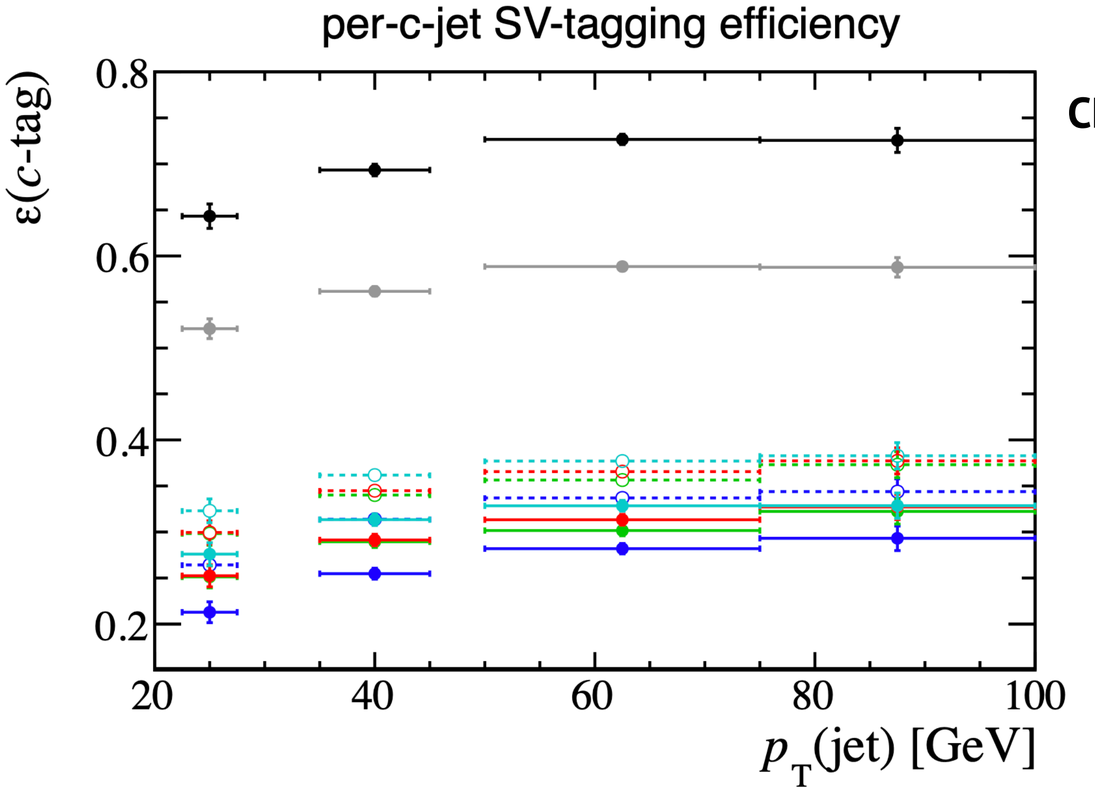
- **The LHCb Collaboration has proposed an upgrade for the HL-LHC era.**
- The goal is to collect **300 fb$^{-1}$** of integrated luminosity with an **improved detector**.



# Prospects on H→ cc̄ search at HL-LHC

- With the **new Vertex Locator** the impact parameter resolution is expected to improve.
- An optimized selection for SV from c-hadron decays can be applied.
- **In our estimation the c-jet SV-tagging efficiency improves roughly from 25% to 35%.**




CERN-2019-007

Perfect detector, i.e. has true SV in kinematic fiducial region.

Perfect IP resolution, but including RECO efficiency (assumed to be same as Run 1, which may not be true), etc.

Phase-II Scenario 2
Phase-II Scenario 1
Run 3
Run1
Solid: IP $X^2 > 16$ (as in Run 1)
Dashed: IP $X^2 > 9$

# Extrapolation of H→ cc̄ search to 300 fb$^{-1}$ at 14 TeV

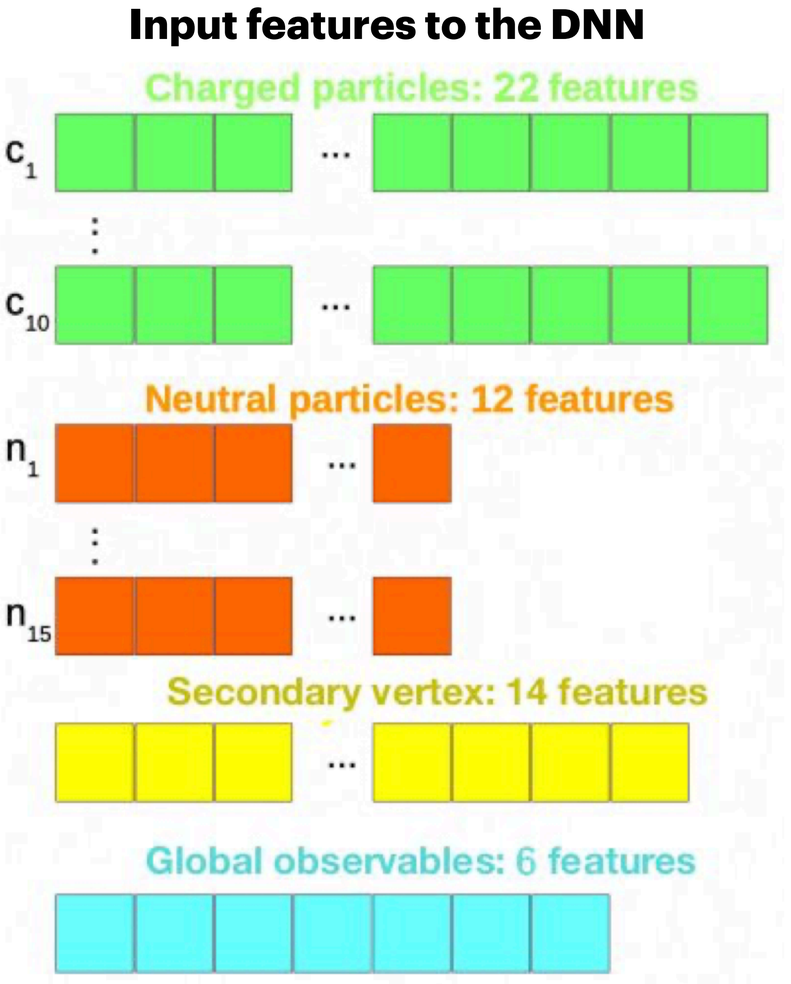
- The **starting point** is the upper limit on the cross section set with 8 TeV data: **6400 x SM**.
- Step-by-step we have the following improvements:
    - **higher cross section and acceptance at 14 TeV and 300 fb$^{-1}$ → 50 x SM**;
    - **improved SV-tagging efficiency → 13x SM**;
    - **improved electron reconstruction → 9 x SM**;
    - **improved background removal → 6 x SM**.

CERN-2019-007

- A **toy study with simulation** leaded to a similar result → **5-10 x SM**.
- This translates to an upper limit of **$y^c < 2\text{-}3\ y^c_{SM}$** on the Hcc coupling.
- **The ATLAS/CMS extrapolation with 2 x 3000 fb$^{-1}$** leads to **$y^c < 6\ y^c_{SM}$** with the optimistic c-tagging performance: **we are in the same game of the General Purpose Detectors**!

# Further improvements on tagging

- The flavour tagging algorithm presented in the previous slides relies on the SV reconstruction.
- We are working on an advanced algorithm that takes in input the complete **jet substructure, even if SV is not found in the jet.**
- Then a machine learning method (**Deep Neural Network**) is used to obtain the probability for b-, c- and light jets.
- Then the working point can be chosen by cutting on these 3 probabilities, in order to obtain the best significance for $H \rightarrow c\bar{c}$.

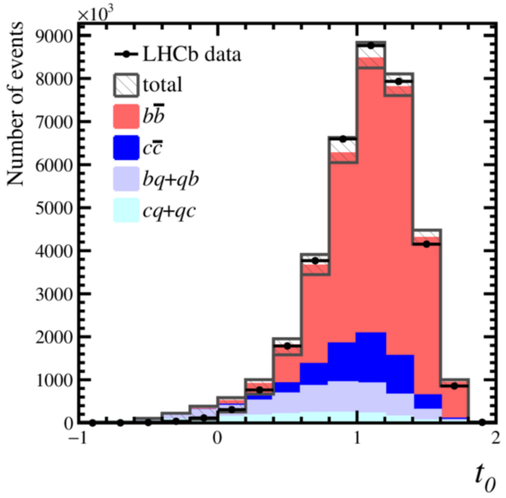


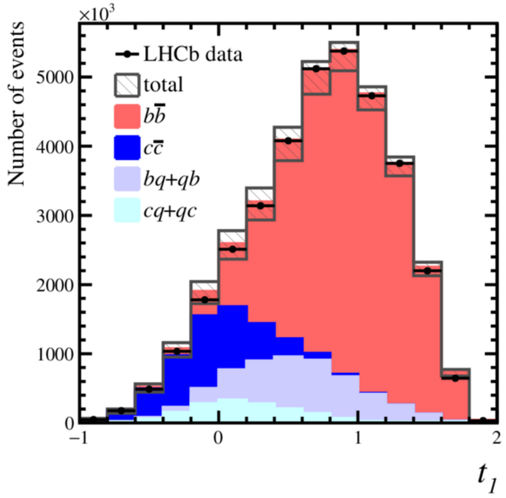
# Inclusive Higgs search in Run 2

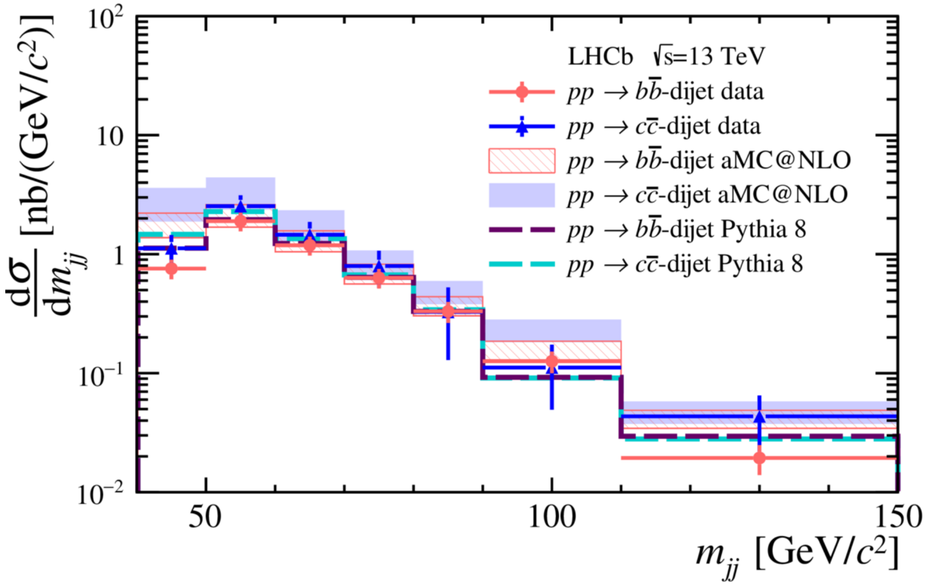
- The **inclusive dataset with two SV-tagged jets (1.6 fb$^{-1}$ at 13 TeV)** has been analyzed to measure the **differential $b\bar{b}$ and $c\bar{c}$ cross section**.
- The technique for disentangle the **$b\bar{b}$** and **$c\bar{c}$** processes has been demonstrated, and the the measured cross sections are compatible with the expectations.
- This is an important step for the **search for inclusive H→$b\bar{b}$ and H→$c\bar{c}$ (so mainly produced by gg fusion)**: **the analysis is on-going**!

$$t_0 = \mathrm{BDT}_{bc|q}(j_0) + \mathrm{BDT}_{bc|q}(j_1),$$
$$t_1 = \mathrm{BDT}_{b|c}(j_0) + \mathrm{BDT}_{b|c}(j_1).$$

**JHEP 02 (2021) 023**



# Conclusions

- **Higgs physics at LHCb is possible!**
- LHCb has already developed techniques to search for **W/Z + H(→cc̄)** by studying Run 1 data.
- **In the HL-LHC era, LHCb could push the limit on the Higgs-charm coupling to 2-3 times the SM expectation**.
- Further improvements are expected: **new tagging methods, the search in the inclusive channel** (without associated production).
- **At least for now the H→cc̄ observation is excluded even at the HL-LHC, but we won't be so far**.

# Thanks for your attention!

# Backup